

(1) GENERAL INFORMATION:

- (i) APPLICANT: ZEBEDEE, SUZANNE  
INCHAUPEPE, GENEVIEVE  
NASOFF, MARC S.  
PRINCE, ALFRED M.  
HELTING, TORSTEN B.  
NUNN, MICHAEL F.
- (ii) TITLE OF INVENTION: METHODS AND SYSTEMS FOR PRODUCING  
RECOMBINANT VIRAL ANTIGENS
- (iii) NUMBER OF SEQUENCES: 29
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Kenyon & Kenyon
  - (B) STREET: One Broadway
  - (C) CITY: New York
  - (D) STATE: NY
  - (E) COUNTRY: USA
  - (F) ZIP: 10004
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy Disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: WordPerfect 6.1 (ASCII text)
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: not yet assigned
  - (B) FILING DATE: concurrent herewith
  - (C) CLASSIFICATION: not yet assigned
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US08/563,733
  - (B) FILING DATE: 28-NOV-1995

- (A) APPLICATION NUMBER: US08/049,531
- (B) FILING DATE: 20-APR-1993
  
- (A) APPLICATION NUMBER: US07/344,237
- (B) FILING DATE: 26-APR-1989
  
- (A) APPLICATION NUMBER: US07/191,229
- (B) FILING DATE: 06-MAY-1988
  
- (A) APPLICATION NUMBER: US07/206,499
- (B) FILING DATE: 13-JUN-1988
  
- (A) APPLICATION NUMBER: US07/258,016
- (B) FILING DATE: 14-OCT-1988
  
- (A) APPLICATION NUMBER: US08/272,271
- (B) FILING DATE: 08-JUL-1994
  
- (A) APPLICATION NUMBER: US07/616,369
- (B) FILING DATE: 21-NOV-1990
  
- (A) APPLICATION NUMBER: US07/573,643
- (B) FILING DATE: 27-AUG-1990

(viii) ATTORNEY/AGENT INFORMATION (O):

- (A) NAME: M. Lisa Wilson, Esq.
- (B) REGISTRATION NUMBER: 34,045
- (C) REFERENCE/DOCKET NUMBER: 55467/69

(ix) TELECOMMUNICATION INFORMATION (O)

- (A) TELEPHONE: (212)908-6366
- (B) TELEFAX: (212)425-5288
- (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 795 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16-789

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AGGAGGGTTT TTCAT ATG CCA ATC GTG CAG AAC ATC CAG GGG CAA ATG GTA 51
      Met Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val
                        5                      10

CAT CAG GCC ATA TCA CCT AGA ACT TTA AAT GCA TGG GTA AAA GTA GTA 99
His Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val
      15                      20                      25

GAA GAG AAG GCT TTC AGC CCA GAA GTG ATA CCC ATG TTT TCA GCA TTA 147
Glu Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu
      30                      35                      40

TCA GAA GGA GCC ACC CCA CAA GAT TTA AAC ACC ATG CTA AAC ACA GTG 195
Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val
      45                      50                      55                      60
```

GGG GGA CAT CAA GCA GCC ATG CAA ATG TTA AAA GAG ACC ATC AAT GAG	243
Gly Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu	
65 70 75	
GAA GCT GCA GAA TGG GAT AGA GTG CAT CCA GTG CAT GCA GGG CCT ATT	291
Glu Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile	
80 85 90	
GCA CCA GGC CAG ATG AGA GAA CCA AGG GGA AGT GAC ATA GCA GGA ACT	339
Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr	
95 100 105	
ACT AGT ACC CTT CAG GAA CAA ATA GGA TGG ATG ACA AAT AAT CCA CCT	387
Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro	
110 115 120	
ATC CCA GTA GGA GAA ATT TAT AAA AGA TGG ATA ATC CTG GGA TTA AAT	435
Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn	
125 130 135 140	
AAA ATA GTA AGA ATG TAT AGC CCT ACC AGC ATT CTG GAC ATA AGA CAA	483
Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln	
145 150 155	
GGA CCA AAG GAA CCC TTT AGA GAC TAT GTA GAC CGG TTC TAT AAA ACT	531
Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr	
160 165 170	
CTA AGA GCC GAG CAA GCT TCA CAG GAG GTA AAA AAT TGG ATG ACA GAA	579
Leu Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu	
175 180 185	
ACC TTG TTG GTC CAA AAT GCG AAC CCA GAT TGT AAG ACT ATT TTA AAA	627
Thr Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys	
190 195 200	

GCA TTG GGA CCA GCG GCT ACA CTA GAA GAA ATG ATG ACA GCA TGT CAG 675  
 Ala Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln  
 205 210 215 220

GGA GTA GGA GGA CCC AAA AAT CAA CAA TTA TTA TCC TTA TGG GGG TGT 723  
 Gly Val Gly Gly Pro Lys Asn Gln Gln Leu Leu Ser Leu Trp Gly Cys  
 225 230 235

AAA GGG AAA CTT GTT TGT TAT ACT TCC GTT AAA TGG AAT GGA CCC GGC 771  
 Lys Gly Lys Leu Val Cys Tyr Thr Ser Val Lys Trp Asn Gly Pro Gly  
 240 245 250

CAT AAG GCA AGA GTT TTG TAA TAA 795  
 His Lys Ala Arg Val Leu  
 255

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 258 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala Ile  
 5 10 15

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala  
 20 25 30

Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala  
 35 40 45

Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	Gly	His	Gln
50						55					60				
Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu	Ala	Ala	Glu
65					70					75					80
Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala	Pro	Gly	Gln
			85						90					95	
Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	Ser	Thr	Leu
		100						105					110		
Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro	Pro	Ile	Pro	Val	Gly
		115					120					125			
Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	Ile	Val	Arg
	130					135					140				
Met	Tyr	Ser	Pro	Thr	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly	Pro	Lys	Glu
145					150					155					160
Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Tyr	Lys	Thr	Leu	Arg	Ala	Glu
			165						170					175	
Gln	Ala	Ser	Gln	Glu	Val	Lys	Asn	Trp	Met	Thr	Glu	Thr	Leu	Leu	Val
		180						185					190		
Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Lys	Ala	Leu	Gly	Pro
		195					200					205			
Ala	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly	Val	Gly	Gly
	210					215					220				
Pro	Lys	Asn	Gln	Gln	Leu	Leu	Ser	Leu	Trp	Gly	Cys	Lys	Gly	Lys	Leu
225					230					235					240

Val Cys Tyr Thr Ser Val Lys Trp Asn Gly Pro Gly His Lys Ala Arg  
245 250 255

Val Leu

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 795 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16-789

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGGAGGGTTT TTCAT ATG CCA ATC GTG CAG AAC ATC CAG GGG CAA ATG GTA	51
Met Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val	
5 10	
CAT CAG GCC ATA TCA CCT AGA ACT TTA AAT GCA TGG GTA AAA GTA GTA	99
His Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val	
15 20 25	
GAA GAG AAG GCT TTC AGC CCA GAA GTG ATA CCC ATG TTT TCA GCA TTA	147
Glu Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu	
30 35 40	

TCA GAA GGA GCC ACC CCA CAA GAT TTA AAC ACC ATG CTA AAC ACA GTG	195
Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val	
45 50 55 60	
GGG GGA CAT CAA GCA GCC ATG CAA ATG TTA AAA GAG ACC ATC AAT GAG	243
Gly Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu	
65 70 75	
GAA GCT GCA GAA TGG GAT AGA GTG CAT CCA GTG CAT GCA GGG CCT ATT	291
Glu Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile	
80 85 90	
GCA CCA GGC CAG ATG AGA GAA CCA AGG GGA AGT GAC ATA GCA GGA ACT	339
Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr	
95 100 105	
ACT AGT ACC CTT CAG GAA CAA ATA GGA TGG ATG ACA AAT AAT CCA CCT	387
Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro	
110 115 120	
ATC CCA GTA GGA GAA ATT TAT AAA AGA TGG ATA ATC CTG GGA TTA AAT	435
Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn	
125 130 135 140	
AAA ATA GTA AGA ATG TAT AGC CCT ACC AGC ATT CTG GAC ATA AGA CAA	483
Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln	
145 150 155	
GGA CCA AAG GAA CCC TTT AGA GAC TAT GTA GAC CGG TTC TAT AAA ACT	531
Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr	
160 165 170	
CTA AGA GCC GAG CAA GCT TCA CAG GAG GTA AAA AAT TGG ATG ACA GAA	579
Leu Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu	
175 180 185	



ACC TTG TTG GTC CAA AAT GCG AAC CCA GAT TGT AAG ACT ATT TTA AAA 627  
 Thr Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys  
 190 195 200

GCA TTG GGA CCA GCG GCT ACA CTA GAA GAA ATG ATG ACA GCA TGT CAG 675  
 Ala Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln  
 205 210 215 220

GGA GTA GGA GGA CCC AAA AAT CAA CAA AGA TTA AAT TTA TGG GGG TGT 723  
 Gly Val Gly Gly Pro Lys Asn Gln Gln Arg Leu Asn Leu Trp Gly Cys  
 225 230 235

AAA GGG AAA CTT ATT TGT TAT ACT TCC GTT AAA TGG AAT GGA CCC GGC 771  
 Lys Gly Lys Leu Ile Cys Tyr Thr Ser Val Lys Trp Asn Gly Pro Gly  
 240 245 250

CAT AAG GCA AGA GTT TTG TAA TAA 795  
 His Lys Ala Arg Val Leu

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 258 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala Ile  
 5 10 15

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala  
 20 25 30

Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala	Leu	Ser	Glu	Gly	Ala	35	40	45
Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	Gly	His	Gln	50	55	60
Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu	Ala	Ala	Glu	65	70	75 80
Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala	Pro	Gly	Gln	85	90	95
Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	Ser	Thr	Leu	100	105	110
Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro	Pro	Ile	Pro	Val	Gly	115	120	125
Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	Ile	Val	Arg	130	135	140
Met	Tyr	Ser	Pro	Thr	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly	Pro	Lys	Glu	145	150	155 160
Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Tyr	Lys	Thr	Leu	Arg	Ala	Glu	165	170	175
Gln	Ala	Ser	Gln	Glu	Val	Lys	Asn	Trp	Met	Thr	Glu	Thr	Leu	Leu	Val	180	185	190
Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Lys	Ala	Leu	Gly	Pro	195	200	205
Ala	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly	Val	Gly	Gly	210	215	220

Pro Lys Asn Gln Gln Arg Leu Asn Leu Trp Gly Cys Lys Gly Lys Leu  
 225 230 235 240

Ile Cys Tyr Thr Ser Val Lys Trp Asn Gly Pro Gly His Lys Ala Arg  
 245 250 255

Val Leu

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 795 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16-789

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGGAGGGTTT TTCAT ATG CCA ATC GTG CAG AAC ATC CAG GGG CAA ATG GTA 51  
 Met Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val  
 5 10

CAT CAG GCC ATA TCA CCT AGA ACT TTA AAT GCA TGG GTA AAA GTA GTA 99  
 His Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val  
 15 20 25

GAA GAG AAG GCT TTC AGC CCA GAA GTG ATA CCC ATG TTT TCA GCA TTA	147
Glu Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu	
30 35 40	
TCA GAA GGA GCC ACC CCA CAA GAT TTA AAC ACC ATG CTA AAC ACA GTG	195
Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val	
45 50 55 60	
GGG GGA CAT CAA GCA GCC ATG CAA ATG TTA AAA GAG ACC ATC AAT GAG	243
Gly Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu	
65 70 75	
GAA GCT GCA GAA TGG GAT AGA GTG CAT CCA GTG CAT GCA GGG CCT ATT	291
Glu Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile	
80 85 90	
GCA CCA GGC CAG ATG AGA GAA CCA AGG GGA AGT GAC ATA GCA GGA ACT	339
Ala Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr	
95 100 105	
ACT AGT ACC CTT CAG GAA CAA ATA GGA TGG ATG ACA AAT AAT CCA CCT	387
Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro	
110 115 120	
ATC CCA GTA GGA GAA ATT TAT AAA AGA TGG ATA ATC CTG GGA TTA AAT	435
Ile Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn	
125 130 135 140	
AAA ATA GTA AGA ATG TAT AGC CCT ACC AGC ATT CTG GAC ATA AGA CAA	483
Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln	
145 150 155	
GGA CCA AAG GAA CCC TTT AGA GAC TAT GTA GAC CGG TTC TAT AAA ACT	531
Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr	
160 165 170	

CTA AGA GCC GAG CAA GCT TCA CAG GAG GTA AAA AAT TGG ATG ACA GAA	579
Leu Arg Ala Glu Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu	
175 180 185	
ACC TTG TTG GTC CAA AAT GCG AAC CCA GAT TGT AAG ACT ATT TTA AAA	627
Thr Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys	
190 195 200	
GCA TTG GGA CCA GCG GCT ACA CTA GAA GAA ATG ATG ACA GCA TGT CAG	675
Ala Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln	
205 210 215 220	
GGA GTA GGA GGA CCA CAA AAT CAA CAA CTT TTA AAT TTA TGG GGG TGT	723
Gly Val Gly Gly Pro Gln Asn Gln Gln Leu Leu Asn Leu Trp Gly Cys	
225 230 235	
AGA GGG AAA GCT ATT TGT TAT ACT TCC GTT CAA TGG AAT GGA CCC GGC	771
Arg Gly Lys Ala Ile Cys Tyr Thr Ser Val Gln Trp Asn Gly Pro Gly	
240 245 250	
CAT AAG GCA AGA GTT TTG TAA TAA	795
His Lys Ala Arg Val Leu	
255	

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 258 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala Ile

	5		10		15										
Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Val	Glu	Glu	Lys	Ala
	20						25					30			
Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala	Leu	Ser	Glu	Gly	Ala
	35						40					45			
Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	Gly	His	Gln
	50					55					60				
Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu	Ala	Ala	Glu
65				70						75					80
Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala	Pro	Gly	Gln
			85						90					95	
Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	Ser	Thr	Leu
			100					105						110	
Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro	Pro	Ile	Pro	Val	Gly
		115					120					125			
Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	Ile	Val	Arg
	130					135						140			
Met	Tyr	Ser	Pro	Thr	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly	Pro	Lys	Glu
145					150					155					160
Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Tyr	Lys	Thr	Leu	Arg	Ala	Glu
				165					170					175	
Gln	Ala	Ser	Gln	Glu	Val	Lys	Asn	Trp	Met	Thr	Glu	Thr	Leu	Leu	Val
		180						185					190		
Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Lys	Ala	Leu	Gly	Pro

195

200

205

Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly  
210 215 220

Pro Gln Asn Gln Gln Leu Leu Asn Leu Trp Gly Cys Arg Gly Lys Ala  
225 230 235 240

Ile Cys Tyr Thr Ser Val Gln Trp Asn Gly Pro Gly His Lys Ala Arg  
245 250 255

Val Leu

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 378 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 16-375

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGGAGGGTTT TTCAT ATG AGC ACG ATT CCC AAA CCT CAA AGA AAA ACC AAA 51  
Met Ser Thr Ile Pro Lys Pro Gln Arg Lys Thr Lys  
5 10

CGT AAC ACC AAC CGT CGC CCA CAG GAC GTC AAG TTC CCG GGT GGC GGT	99
Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly	
15 20 25	

CAG ATC GTT GGT GGA GTT TAC TTG TTG CCG CGC AGG GGC CCT AGA TTG	147
Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu	
30 35 40	

GGT GTG CGC GCG ACG AGG AAG ACT TCC GAG CGG TCG CAA CCT CGA GGT	195
Gly Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly	
45 50 55 60	

AGA CGT CAG CCT ATC CCC AAG GTG CGT CGG CCG GAG GGC AGG ACC TGG	243
Arg Arg Gln Pro Ile Pro Lys Val Arg Arg Pro Glu Gly Arg Thr Trp	
65 70 75	

GCT CAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAT GAG GGT TGC GGG	291
Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly	
80 85 90	

TGG GCG GGA TGG CTC CTG TCT CCC CGT GGC TCT CGG CCT AGC TGG GGC	339
Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly	
95 100 105	

CCC ACA GAC CCC CGG CGT AGG TCG CGC AAT TTG GGT TAA	378
Pro Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly	
110 115 120	

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Ser Thr Ile Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
5 10 15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly  
20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
35 40 45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
50 55 60

Ile Pro Lys Val Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly  
65 70 75 80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp  
85 90 95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro  
100 105 110

Arg Arg Arg Ser Arg Asn Leu Gly  
115 120

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 378 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 16-375

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGAGGGTTT TTCAT ATG AGC ACG ATT CCC AAA CCT CAA AGA AAA ACC AAA	51
Met Ser Thr Ile. Pro Lys Pro Gln Arg Lys Thr Lys	
5 10	
CGT AAC ACC AAC CGT CGC CCA CAG GAC GTC AAG TTC CCG GGT GGC GGT	99
Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly	
15 20 25	
CAG ATC GTT GGT GGA GTT TAC TTG TTG CCG CGC AGG GGC CCT AGA TTG	147
Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu	
30 35 40	
GGT GTG CGC GCG ACG AGG AAG ACT TCC GAG CGG TCG CAA CCT CGA GGT	195
Gly Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly	
45 50 55 60	
AGA CGT CAG CCT ATC CCC AAG GCA CGT CGG CCC GAG GGC AGG ACC TGG	243
Arg Arg Gln Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp	
65 70 75	
GCT CAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAT GAG GGT TGC GGG	291
Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly	
80 85 90	
TGG GCG GGA TGG CTC CTG TCT CCC CGT GGC TCT CGG CCT AGC TGG GGC	339
Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly	

95

100

105

CCC ACA GAC CCC CGG CGT AGG TCG CGC AAT TTG GGT TAA 378  
 Pro Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly  
 110 115 120

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear.

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ser Thr Ile Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
                   5                  10                  15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly  
                   20                  25                  30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
                   35                  40                  45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
                   50                  55                  60

Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly  
                   65                  70                  75                  80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp  
                   85                  90                  95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro

100

105

110

Arg Arg Arg Ser Arg Asn Leu Gly  
115 120

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 378 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear.

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16-375

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGGAGGGTTT TTCAT ATG AGC ACG ATT CCC AAA CCT CAA AGA AAA ACC AAA	51
Met Ser Thr Ile Pro Lys Pro Gln Arg Lys Thr Lys	
5 10	
CGT AAC ACC AAC CGT CGC CCA CAG GAC GTC AAG TTC CCG GGT GGC GGT	99
Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly	
15 20 25	
CAG ATC GTT GGT GGA GTT TAC TTG TTG CCG CGC AGG GGC CCT AGA TTG	147
Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu	
30 35 40	

GGT GTG CGC GCG ACG AGG AAG ACT TCC GAG CGG TCG CAA CCT CGA GGT 195  
 Gly Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly  
 45 50 55 60

AGA CGT CAG CCT ATC CCC AAG GAC CGT CGG TCC ACG GGC AAG TCC TGG 243  
 Arg Arg Gln Pro Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp  
 65 70 75

GGT AAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAT GAG GGT TGC GGG 291  
 Gly Lys Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly  
 80 85 90

TGG GCG GGA TGG CTC CTG TCT CCC CGT GGC TCT CGG CCT AGC TGG GGC 339  
 Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly  
 95 100 105

CCC ACA GAC CCC CGG CGT AGG TCG CGC AAT TTG GGT TAA 378  
 Pro Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly  
 110 115 120

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Ser Thr Ile Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
 5 10 15  
 Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly  
 20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
35 40 45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
50 55 60

Ile Pro Lys Asp Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly  
65 70 75 80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp  
85 90 95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro  
100 105 110

Arg Arg Arg Ser Arg Asn Leu Gly  
115 120

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 378 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16-375

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGGAGGGTTT TTCAT ATG AGC ACG ATT CCC AAA CCT CAA AGA AAA ACC AAA	51
Met Ser Thr Ile Pro Lys Pro Gln Arg Lys Thr Lys	
5 10	
CGT AAC ACC AAC CGT CGC CCA CAG GAC GTC AAG TTC CCG GGT GGC GGT	99
Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly	
15 20 25	
CAG ATC GTT GGT GGA GTT TAC TTG TTG CCG CGC AGG GGC CCT AGA TTG	147
Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu	
30 35 40	
GGT GTG CGC GCG ACG AGG AAG ACT TCC GAG CGG TCG CAA CCT CGA GGT	195
Gly Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly	
45 50 55 60	
AGA CGT CAG CCT ATC CCC AAG GCA CGT CGG TCC GAG GGC AGG TCC TGG	243
Arg Arg Gln Pro Ile Pro Lys Ala Arg Arg Ser Glu Gly Arg Ser Trp	
65 70 75	
GCT CAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAT GAG GGT TGC GGG	291
Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly	
80 85 90	
TGG GCG GGA TGG CTC CTG TCT CCC CGT GGC TCT CGG CCT AGC TGG GGC	339
Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly	
95 100 105	
CCC ACA GAC CCC CGG CGT AGG TCG CGC AAT TTG GGT TAA	378
Pro Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly	
110 115 120	

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ser Thr Ile Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
5 10 15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly  
20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
35 40 45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
50 55 60

Ile Pro Lys Ala Arg Arg Ser Glu Gly Arg Ser Trp Ala Gln Pro Gly  
65 70 75 80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp  
85 90 95

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro  
100 105 110

Arg Arg Arg Ser Arg Asn Leu Gly  
115 120

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:



- (A) LENGTH: 381 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16-378

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AGGAGGGTTT TTCAT ATG CCT ATT CAT CAT CAT CAT CAT CAT GGC CCG GGC	51
Met Pro Ile His His His His His His His Gly Pro Gly	
5 10	
TCC GTC ACT GTG TCC CAT CCT AAC ATC GAG GAG GTT GCT CTG TCC ACC	99
Ser Val Thr Val Ser His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr	
15 20 25	
ACC GGA GAG ATC CCC TTT TAC GGC AAG GCT ATC CCC CTC GAG GTG ATC	147
Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile	
30 35 40	
AAG GGG GGA AGA CAT CTC ATC TTC TGC CAC TCA AAG AAG AAG TGC GAC	195
Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp	
45 50 55 60	
GAG CTC GCC GCG AAG CTG GTC GCA TTG GGC ATC AAT GCC GTG GCC TAC	243
Glu Leu Ala Ala Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala Tyr	
65 70 75	

TAC CGC GGT CTT GAC GTG TCT GTC ATC CCG ACC AGC GGC GAT GTT GTC 291  
 Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val  
                     80                                    85                                    90

GTC GTG TCA ACC GAT GCT CTC ATG ACT GGC TTT ACC GGC GAC TTC GAC 339  
 Val Val Ser Thr Asp Ala Leu Met Thr Gly Phe Thr Gly Asp Phe Asp  
                     95                                    100                                    105

TCG GTG ATA GAC TGC AAT ACG GGT ACC GAG CTC GAA TTC TAA 381  
 Ser Val Ile Asp Cys Asn Thr Gly Thr Glu Leu Glu Phe  
                     110                                    115                                    120

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 121 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Pro Ile His His His His His His Gly Pro Gly Ser Val Thr Val  
                     5                                    10                                    15

Ser His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile  
                     20                                    25                                    30

Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg  
                     35                                    40                                    45

His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala  
                     50                                    55                                    60

Lys Leu Val Ala Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu

65		70		75		80									
Asp	Val	Ser	Val	Ile	Pro	Thr	Ser	Gly	Asp	Val	Val	Val	Val	Ser	Thr
				85					90					95	
Asp	Ala	Leu	Met	Thr	Gly	Phe	Thr	Gly	Asp	Phe	Asp	Ser	Val	Ile	Asp
				100					105					110	
Cys	Asn	Thr	Gly	Thr	Glu	Leu	Glu	Phe							
				115										120	

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 774 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 16-771

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGAGGGTTT	TTCAT	ATG	TCC	CCT	ATT	CTA	GGT	TAT	TGG	AAA	ATT	AAG	GGC	51		
		Met	Ser	Pro	Ile	Leu	Gly	Tyr	Trp	Lys	Ile	Lys	Gly			
						5						10				
CTT	GTG	CAA	CCC	ACT	CGA	CTT	CTT	TTG	GAA	TAT	CTT	GAA	GAA	AAA	TAT	99

Leu Val Gln Pro Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr	
15 20 25	
GAA GAG CAT TTG TAT GAG CGC GAT GAA GGT GAT AAA TGG CGA AAC AAA	147
Glu Glu His Leu Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys	
30 35 40	
AAG TTT GAA TTG GGT TTG GAG TTT CCC AAT CTT CCT TAT TAT ATT GAT	195
Lys Phe Glu Leu Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp	
45 50 55 60	
GGT GAT GTT AAA TTA ACA CAG TCT ATG GCC ATC ATA CGT TAT ATA GCT	243
Gly Asp Val Lys Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala	
65 70 75	
GAC AAG CAC AAC ATG TTG GGT GGT TGT CCA AAA GAG CGT GCA GAG ATT	291
Asp Lys His Asn Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile	
80 85 90	
TCA ATG CTT GAA GGA GCG GTT TTG GAT ATT AGA TAC GGT GTT TCG AGA	339
Ser Met Leu Glu Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg	
95 100 105	
ATT GCA TAT AGT AAA GAC TTT GAA ACT CTC AAA GTT GAT TTT CTT AGC	387
Ile Ala Tyr Ser Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser	
110 115 120	
AAG CTA CCT GAA ATG CTG AAA ATG TTC GAA GAT CGT TTA TGT CAT AAA	435
Lys Leu Pro Glu Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys	
125 130 135 140	
ACA TAT TTA AAT GGT GAT CAT GTA ACC CAT CCT GAC TTC ATG TTG TAT	483
Thr Tyr Leu Asn Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr	
145 150 155	
GAC GCT CTT GAT GTT GTT TTA TAC ATG GAC CCA ATG TGC CTG GAT GCG	531

Asp Ala Leu Asp Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala  
160 165 170

TTC CCA AAA TTA GTT TGT TTT AAA AAA CGT ATT GAA GCT ATC CCA CAA 579  
Phe Pro Lys Leu Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln  
175 180 185

ATT GAT AAG TAC TTG AAA TCC AGC AAG TAT ATA GCA TGG CCT TTG CAG 627  
Ile Asp Lys Tyr Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln  
190 195 200

GGC TGG CAA GCC ACG TTT GGT GGT GGC GAC CAT CCT CCA AAA TCG GAT 675  
Gly Trp Gln Ala Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp  
205 210 215 220

CTG GTT CCG CGT GGA TCC GAC GTC AAG TTC CCG GGT GGC GGT CAG ATC 723  
Leu Val Pro Arg Gly Ser Asp Val Lys Phe Pro Gly Gly Gly Gln Ile  
225 230 235

GTT GGT GGA GTT TAC TTG TTG CCG CGC AGG GAA TTC ATC GTG ACT GAC 771  
Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Glu Phe Ile Val Thr Asp  
240 245 250

TGA 774

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 252 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met	Ser	Pro	Ile	Leu	Gly	Tyr	Trp	Lys	Ile	Lys	Gly	Leu	Val	Gln	Pro			
				5					10					15				
Thr	Arg	Leu	Leu	Leu	Glu	Tyr	Leu	Glu	Glu	Lys	Tyr	Glu	Glu	His	Leu			
			20					25					30					
Tyr	Glu	Arg	Asp	Glu	Gly	Asp	Lys	Trp	Arg	Asn	Lys	Lys	Phe	Glu	Leu			
		35					40					45						
Gly	Leu	Glu	Phe	Pro	Asn	Leu	Pro	Tyr	Tyr	Ile	Asp	Gly	Asp	Val	Lys			
		50				55					60							
Leu	Thr	Gln	Ser	Met	Ala	Ile	Ile	Arg	Tyr	Ile	Ala	Asp	Lys	His	Asn			
65					70					75				80				
Met	Leu	Gly	Gly	Cys	Pro	Lys	Glu	Arg	Ala	Glu	Ile	Ser	Met	Leu	Glu			
				85				90						95				
Gly	Ala	Val	Leu	Asp	Ile	Arg	Tyr	Gly	Val	Ser	Arg	Ile	Ala	Tyr	Ser			
		100						105					110					
Lys	Asp	Phe	Glu	Thr	Leu	Lys	Val	Asp	Phe	Leu	Ser	Lys	Leu	Pro	Glu			
		115					120					125						
Met	Leu	Lys	Met	Phe	Glu	Asp	Arg	Leu	Cys	His	Lys	Thr	Tyr	Leu	Asn			
		130				135					140							
Gly	Asp	His	Val	Thr	His	Pro	Asp	Phe	Met	Leu	Tyr	Asp	Ala	Leu	Asp			
145					150				155					160				
Val	Val	Leu	Tyr	Met	Asp	Pro	Met	Cys	Leu	Asp	Ala	Phe	Pro	Lys	Leu			
				165				170					175					
Val	Cys	Phe	Lys	Lys	Arg	Ile	Glu	Ala	Ile	Pro	Gln	Ile	Asp	Lys	Tyr			
			180					185					190					

Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala  
195 200 205

Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Val Pro Arg  
210 215 220

Gly Ser Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val  
225 230 235 240

Tyr Leu Leu Pro Arg Arg Glu Phe Ile Val Thr Asp  
245 250

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCAAAATTAC CATATGCCAA TCGTGCAGAA C

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GACCCGGCCA TAAGGCAAGA GTTTTGTAAT AAG

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GATCCTTATT ACAAACTCT TGCCTTATGG CCGG

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs

(B) TYPE: nucleic acid



(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCTCGCATAT GAGCACGATT CCCAAACC

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GACGAATTCT TAACCCAAAT TGCGCGACCT AC

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GATCCGACGT CAAGTTCCCG GGTGGCGGTC AGATCGTTGG TGGAGTTTAC TTGTTGCCGC  
GCAGGG

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AATTCCTGC GCGGCAACAA GTAAACTCCA CCAACGATCT GACCGCCACC CGGGAAGTTG  
ACGTCG

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGAATTCCAT ATGTCCCCTA TACTAGGT

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CGGAATTCTC ACCTGCGCGG CAACAA

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 52 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: no

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TATGCCTATT CATCATCATC ATCATCATGG CCCGGGAATT CTAAGTAAGT AG

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 54 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GATCCTACTT ACTTAGAATT CCCGGGCCAT GATGATGATG ATGATGAATA GGCA